УДК 681.3.06 (075.32)
DOI: 10.15827/0236-235X.129.054-060
Дата подачи статьи: 28.07.19
2020. Т. 33. № 1. С. 054–060

Методика решения задачи антиспуфинга по ограниченному количеству фотографий

К.Д. Русаков [1], *младший научный сотрудник, rusakov.msk@yandex.ru*
А.А. Генов [2], *д.т.н., профессор, старший научный сотрудник, kt-mati@mail.ru*
С.Ш. Хиль [3], *к.т.н., доцент, skhill@mail.ru*

[1] *Институт проблем управления им. В.А. Трапезникова Российской академии наук, г. Москва, 117997, Россия*
[2] *Центр визуализации и спутниковых информационных технологий, ФНЦ НИИ системных исследований РАН, г. Москва, 117218, Россия*
[3] *Московский авиационный институт (Национальный исследовательский университет), г. Москва, 125993, Россия*

В настоящее время задача предоставления высокого уровня безопасности мобильных устройств, таких как смартфоны и планшеты, посредством биометрических подходов особо актуальна. В статье предложена методика решения задачи антиспуфинга по ограниченному количеству изображений. Исследуются детекции spoof-атак с использованием распечатанных фотографий и экранов мобильных устройств и мониторов. Показаны актуальность исследования и нерешенность задачи в целом.

Рассмотрена структура типовой системы Liveness, состоящей из источника (камеры), препроцессинга получаемых изображений, детекции лиц, модуля признаков и классификатора. В ходе исследования отмечено, что предобработка получаемых изображений является одной из самых главных частей системы, так как вследствие обширности аугментаций признаки spoof-атак выявляются тяжело.

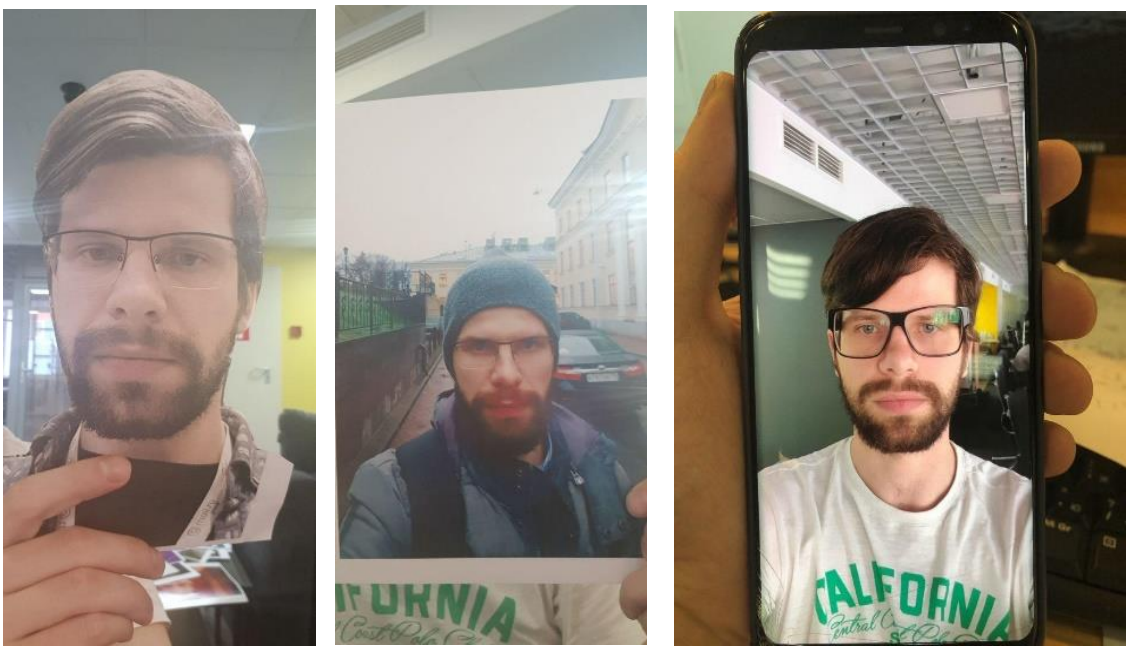
Дается небольшой обзор современных архитектур сверточных нейронных сетей (в терминологии текущей архитектуры – энкодеров), а также показано, что линейный выход сверточных нейронных сетей можно использовать как вход для рекуррентных нейронных сетей типа LSTM. Отмечается, что для детекции лиц наилучшим алгоритмом в условиях текущей архитектуры является MMOD-метод.

Итоговая архитектура системы Liveness представлена как комбинированный подход, состоящий из сверточной нейронной сети, получающей эмбеддинги от каждого кадра, и рекуррентной нейронной сети LSTM, использующей эти эмбеддинги на входе и обучающейся запоминанию последовательности определенных кадров и их характеристик.

Представлены результаты экспериментов, показывающие показатели точности текущих разработок в данной области при условии довольно низких требований к вычислительной мощности. Данная методика позволяет достаточно достоверно определять поддельность фотографии по нескольким кадрам.

***Ключевые слова**: антиспуфинг, живучесть, распознавание лиц, сверточные нейронные сети, рекуррентные нейронные сети.*

Задача предоставления высокого уровня безопасности мобильных устройств, таких как смартфоны и планшеты, посредством биометрических подходов является особо актуальной. В данном направлении активно исследовались системы проверки радужной оболочки [1] и отпечатков пальцев [2], в настоящее время используемые в различных системах безопасности. Эти подходы гарантируют высокую эффективность, однако требуют преднамеренного контакта с устройством, что, по мнению пользователей, неудобно. Для решения этой задачи в качестве альтернативы наибольшую популярность получил метод распознавания лица. Однако он уязвим при различных spoofing-атаках (spoofing – подмена), в которых используются фотографии или видео человека из Интернета или с камеры. В контексте компьютерного зрения spoofing-атака – ситуация, в которой один человек (или программа) успешно маскируется под другого путем фальсификации его визуальных данных, что позволяет получить незаконные преимущества. Например, напечатанные фотографии (рис. 1б), имитирующие маски (рис. 1а) и снимки с экрана (рис. 1в), используются для несанкционированных попыток входа в систему. Кроме того, злоумышленник может захватывать видеопоследовательности, отражающие мимику, непроизвольные изменения лица, например, моргание глаз, и воспроизводить их, чтобы проникнуть в систему безопасности. Некоторые исследователи приложили немало усилий для разработки подходов, позволяющих отли-



а) *б)* *в)*

Рис. 1. Примеры фотографий, использующих spoofing-атаки

Fig. 1. Sample photos using spoofing attacks

чать живые лица от поддельных на основе информации о движении, спектре и качестве изображения.

С одной стороны, чаще всего для противодействия spoofing-атакам используются подходы, основанные на анализе видеопоследовательностей. Они нацелены на выявление естественных изменений лица, например, мигание глаз [3], движение рта [4] и вращение головы [5]. В частности, в [3] авторы детектировали моргание глаз на основе неориентированной условной графической структуры, а в [4] предложили использовать видеопоток части изображения в области рта. Они спроецировали векторы скорости на свою интуитивно понятную модель и извлекли статистику движений губ, чтобы определить, насколько лицо является неподдельным. В [6] предложено использовать корреляции между движениями головы пользователя и фоном, которые указывают на spoofing-атаку. Хотя эти подходы концептуально просты, для отслеживания компонентов лица необходимо достаточно большое количество кадров, что требует увеличения времени обнаружения и дополнительных активных действий пользователя, а также нагружает канал связи в случае удаленного API-сервера.

С другой стороны, методы, основанные на спектре, явно рассчитывают разницу между классами, настоящими и поддельными лицами. В [7] авторы измеряли различия в отражениях между настоящими и поддельными лицами на основе яркости при различном освещении, после чего эти оценочные значения были применены к линейному дискриминанту Фишера. В работе [8] была применена эффективная методика детектирования spoofing-атак всего лишь по единичному экземпляру. Особенность алгоритма заключается в построении карты скорости диффузии к изображениям, пропущенным через специальные фильтры адаптивной нелинейной фильтрации.

Помимо указанных методов, возможно применение эффективных комбинированных алгоритмов, способных улавливать и пространственные, и временные признаки фотографий. Так, специалисты лаборатории компьютерного зрения Мичиганского технического университета в своей работе [9] исследовали два направления борьбы с spoofing-атаками: пространственный анализ (карт глубины лиц) и временной (анализ rPPG-сигнала по видеопоследовательности). Решение о поддельности фотографии принимается совместно по этим двум составляющим.

Несмотря на достаточно большое количество работ в данном направлении, задача противодействия spoofing-атакам до сих пор остается актуальной. Во-первых, ни один из перечисленных подходов не позволяет однозначно

и безошибочно решить задачу, во-вторых, все высокоточные методы имеют большую расчетную сложность и ресурсоемкость, а иногда требуют высокоскоростных каналов передачи данных [10].

Система Liveness

Типичная структурная схема системы обнаружения неподдельного лица (система Liveness) показана на рисунке 2. Чтобы использовать систему Liveness, пользователь должен представить соответствующую биометрическую характеристику сенсору, в данном случае – камере. Захваченные изображения лица предварительно обрабатываются (например, с помощью методов нормализации и удаления шума, пропуска через определенные фильтры и т.п.), поскольку отдельные неподдельные черты лица впоследствии могут быть извлечены в модуле извлечения признаков. Также в модуль предварительной обработки может быть включен модуль детекции лица, целью которого является определение наличия лица на фотографии как такового. Результатом работы модуля извлечения признаков является биометрический шаблон (эмбеддинг), который содержит совокупность элементов (признаков), позволяющих отличить настоящие образцы от поддельных.

Источник (камера). В ходе исследования были изучены различные камеры на различных устройствах. Обычно для подачи входных образцов в систему определения неподдельности лица и его биометрических показателей используются камеры высокого качества на таких телефонах, как Iphone7, SamsungS6, IPad3 и выше. Камеры видимого света являются одними из наиболее используемых устройств, поскольку они дешевле, быстрее, с более высоким разрешением и просты в применении. Однако такие камеры ограничены съемкой только тех изображений, которые имеют видимый световой спектр 3,5–26. Кроме того, камеры плохого качества усложняют определение неподдельного лица, поскольку лица на фото, сделанных с таких камер, выглядят сильно смазанными и зашумленными. Также, например, возможно использование 3D-датчиков, так как они имеют высокую скорость сбора данных, не зависящую от окружающего освещения, с точностью до микронного диапазона. 3D-датчики могут зависеть от вычислений, времени измерения, стоимости и качества, ожидаемых от измерения.

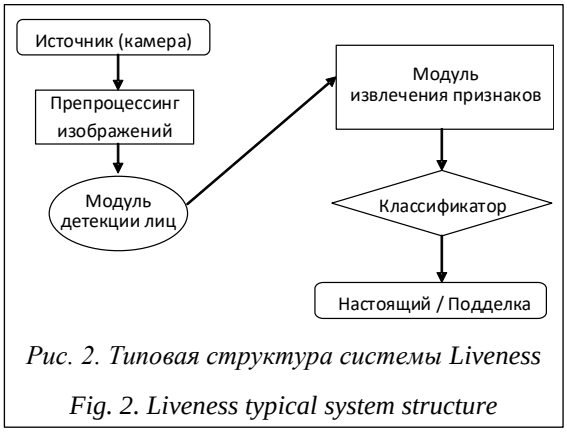


Рис. 2. Типовая структура системы Liveness

Fig. 2. Liveness typical system structure

Препроцессинг. На систему Liveness могут влиять различия в освещении, позе человека и качестве изображения. Чтобы повысить эффективность обнаружения неподдельности фото, несколько систем осуществляют его предварительную обработку, которая обычно включает в себя удаление шума с изображения и пропуск через определенные фильтры (например, баланс белого). После этого, как правило, образцы поступают в модуль детекции лиц для определения факта наличия/отсутствия лица. Далее следуют этапы нормализации для улучшения визуального представления изображений лица для выделения признаков. Методы могут включать сглаживание, размытие, резкость, обнаружение краев или масштабирование. Затем предварительно обработанные образцы направляются в модуль извлечения признаков, чтобы выбрать существенные признаки, позволяющие дифференцировать настоящие образцы от поддельных аналогов.

Детекция лица. Перед определением неподдельности лица необходимо обнаружить его на фото. Большинство методов обнаружения лиц работают с помощью двоичного классификатора, после чего следует шаг так называемого немаксимального подавления, заключающийся в том, что перекрывающиеся прямоугольники, в которых найдено лицо, удаляются. Поскольку число возможных прямоугольников, в которых могут находиться лица, в наборах данных изображений даже умеренного размера чрезвычайно велико, классификатор обычно обучается только на подмножестве прямоугольников. Одним из перспективных и наиболее популярных методов детекции лиц является метод Max-Margin Object Detection (MMOD) [11]. Этот метод оптимизирует все прямоугольники. Он может быть использован для улучшения любого метода обнаружения лиц, который является линейным по изучен-

ным параметрам, таким как HOG (гистограммы направленных градиентов).

Модули извлечения признаков (энкодеры). Для получения важных признаков из изображений принято использовать сверточные нейронные сети (рис. 3). Технически эти сети в процессе обучения и тестирования каждое входное изображение пропускают через серию сверточных слоев с фильтрами, объединениями и полностью связанными слоями.

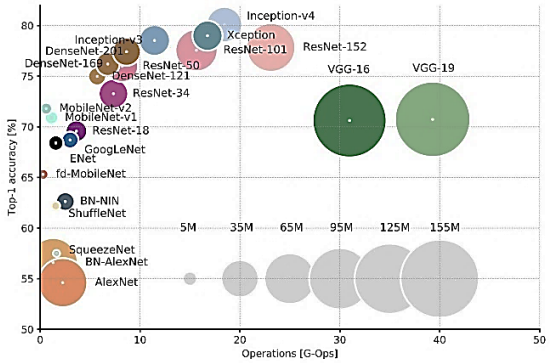


Рис. 3. Наиболее распространенные архитектуры сверточных нейросетей

Fig. 3. The most common convolutional neural network architectures

На выходе последнего линейного слоя у сверточной нейронной сети получается вектор определенной длины, который иногда называется эмбеддингом. Данный вектор содержит наиболее важную информацию о полученном на входе сети изображении. Так как по факту нейросеть сворачивает исходное изображение в небольшой вектор, ее иногда называют энкодером.

Методика детекции spoofing-атак в условиях ограниченного количества фотографий

Основная идея методики детекции spoofing-атак заключается в учете временных характеристик, полученных путем применения наиболее эффективного в этом плане энкодера для объекта съемки.

Как показано на рисунке 4, предлагаемая архитектура позволяет совместно учитывать и признаки, полученные модулем извлечения признаков (CNN), и временные характеристики с помощью сети LSTM. Рекуррентные нейронные сети, основанные на этом подходе, имеют более продвинутый (и более сложный) способ вычисления h^t. Данный способ [12], помимо входных значений и предыдущего состояния сети, использует также фильтры (gates), определяющие, каким образом информация будет использоваться для вычисления как выходных значений на текущем слое y^t, так и значений скрытого слоя на следующем шаге h^{t+1}. Весь процесс вычисления h^t для простоты упоминается как LSTM-слой.

Рассмотрим подробнее структуру LSTM-слоя [13, 14]. Центральным понятием здесь является запоминающий блок (memory cell), который наряду с состоянием сети h вычисляется на каждом шаге, используя текущее входное значение x^t и значение блока на предыдущем шаге c^{t-1}. Входной фильтр (input gate) i^t определяет, насколько значение блока памяти на текущем шаге должно влиять на результат. Значения фильтра варьируются от 0 (полностью игнорировать входные значения) до 1, что обеспечивается областью значений сигмоидальной функции: $i^t = \sigma(W^i x^t + U^i h^{t-1})$.

Фильтр забывания (forget gate) позволяет исключить при вычислениях значения памяти предыдущего шага: $f^t = \sigma(W^f x^t + U^f h^{t-1})$.

На основе всех данных, поступающих в момент времени t, вычисляется состояние блока памяти c^t на текущем шаге, используя фильтры: $c^{*t} = \tanh(Wx^t + Uh^{t-1})$, $f^t = f^t c^{t-1} + i^t c^{*t}$.

Выходной фильтр (output gate) аналогичен двум предыдущим и имеет вид $o^t = \sigma(W^o x^t + U^o h^{t-1})$.

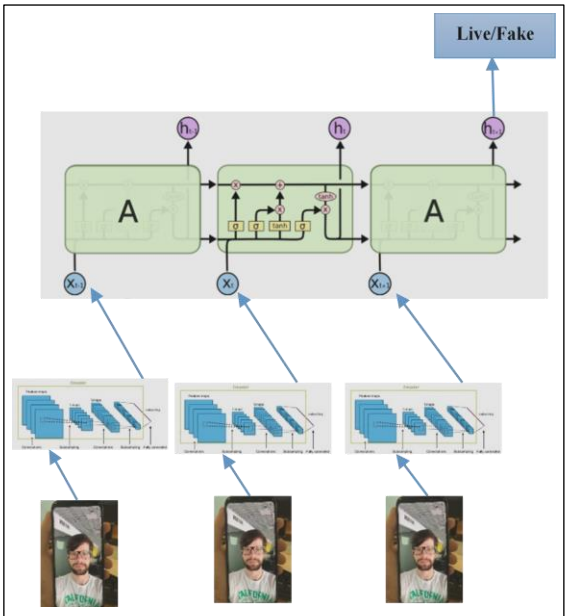


Рис. 4. Архитектура совместного использования CNN и LSTM

Fig. 4. CNN and LSTM sharing architecture

Результаты тестирования

Test results

Тип spoofing-атаки	Количество субъектов	Количество примеров	Точность, %
Фото высокого качества, снятые на высококачественную камеру IPhone8	100	36 000	85
Фото высокого качества на экране монитора	50	15 000	99
Фото высокого качества на экране планшета IPad	30	10 000	85
Видео высокого качества, снятое на IPhone8	30	10 000	91

Итоговое значение LSTM-слоя определяется выходным фильтром (output gate) и нелинейной трансформацией над состоянием блока памяти: $h^t = o^t \tanh(c^t)$.

В качестве энкодера была выбрана стандартная структура ResNet101 с небольшими изменениями линейного слоя. На входе LSTM получены эмбеддинги x^{t+i} и значения целевой переменной (Live/Spoof), которые обучали архитектуру CNN–LSTM, используя бинарную кросс-энтропию как функцию потерь: $H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log(p(y_i)) + (1-y_i)\cdot\log(1-p(y_i))$, где y_i – метка (Live/Spoof); $p(y_i)$ – вероятность того, что метка будет равна 1 для всех N примеров.

Результаты эксперимента

Архитектура тестировалась на данных, полученных при различном освещении с различных мобильных устройств. В таблице приведены итоги тестирования.

Различные типы spoofing-атак были применены к разному количеству уникальных субъектов (уникальных лиц), при этом количество примеров означает все возможные комбинации кадров в соответствии с указанными требованиями в рамках одного уникального субъекта. Точность на тестовой выборке определялась по формуле $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$, где TP – истинно-положительные примеры; TN – истинно-отрицательные примеры; FP – ложно-положительные примеры; FN – ложно-отрицательные примеры.

Заключение

Разработанная методика построена на ранее проведенных исследованиях и тестировании моделей и подходов различных структур в условиях, наиболее приближенных к повседневным. Данная методика позволяет на достаточно высоком уровне определять поддельность фотографии по нескольким кадрам. Среди недостатков модели отмечается все еще присутствующая уязвимость фото высокой четкости.

Публикация выполнена в рамках госзадания ФГУ ФНЦ НИИСИ РАН (выполнение фундаментальных научных исследований ГП 14), тема № 0065-2019-0001 (АААА-А19-119011790077-1).

Литература

1. Bhaganagare B.B., Harale A.D. Iris as biometrics for security system. Proc. ICECCT, Coimbatore, 2017, pp. 1–7. DOI: 10.1109/ICECCT.2017.8117952.

2. Ivanov V.I., Baras J.S. Authentication of fingerprint scanners. Proc. IEEE ICASSP, Prague, 2011, pp. 1912–1915.

3. Pan G., Sun L., Wu Z., Lao S. Eyeblink-based antispoofing in face recognition from a generic Webcamera. Proc. IEEE 11th ICCV'07, Rio de Janeiro, Brazil, 2007, pp. 14–20. DOI: 10.1109/ICCV.2007.4409068.

4. Kollreider K., Fronthaler H., Faraj M.I., Bigun J. Real-time face detection and motion analysis with application in 'liveness' assessment. Proc. IEEE Trans. Inf. Forensics Security, 2007, vol. 2, no. 3, pp. 548–558. DOI: 10.1109/TIFS.2007.902037.

5. Sun L., Pan G., Wu Z., Lao S. Blinking-based live face detection using conditional random fields. Proc. Intern. Conf. Adv. Biometrics, Seoul, Korea, 2007, pp. 252–260.

6. Anjos A., Chakka M.M., Marcel S. Motion-based counter-measures to photo attacks in face recognition. IET Biometrics, 2014, vol. 3, no. 3, pp. 147–158.

7. Charnes A., Cooper W.W. and Rhodes E. Measuring the efficiency of decision making units. Europ. J. Operation Research, 1978, vol. 2, pp. 429–444.

8. Kim Y., Na J., Yoon S., Yi J. Masked fake face detection using radiance measurements. J. Opt. Soc. Amer. A., 2009, vol. 26, no. 4, pp. 760–766.

9. Kim W., Suh S., Han J. Face liveness detection from a single image via diffusion speed model. Proc. IEEE Trans Image Process., 2015, vol. 24, no. 8, pp. 2456–2465. DOI: 10.1109/TIP.2015.2422574.

10. Ефремов И.А., Мамросенко К.А., Решетников В.Н. Методы разработки драйверов графической подсистемы // Программные продукты и системы. 2018. № 3. С. 425–429. DOI: 10.15827/0236-235X.123.425-429.

11. Liu Y., Jourabloo A., Liu X. Learning deep models for face anti-spoofing: Binary or auxiliary supervision. Proc. CVPR, IEEE, 2018. URL: https://ieeexplore.ieee.org/document/8578146/authors#authors (дата обращения: 20.07.2019). DOI: 10.1109/CVPR.2018.00048.

12. King D.E. Max-Margin Object Detection. 2015. URL: https://arxiv.org/abs/1502.00046 (дата обращения: 20.07.2019).

13. Hochreiter S., Schmidhuber J. Long short-term memory. Neural Computation. 1997, vol. 9, no. 8, pp. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.

14. Будыльский Д.В. GRU и LSTM: современные рекуррентные нейронные сети // Молодой ученый. 2015. № 15. С. 51–54.

Software & Systems Received 28.07.19
DOI: 10.15827/0236-235X.129.054-060 2020, vol. 33, no. 1, pp. 054–060

An anti-spoofing methodology for a limited number of photos

K.D. Rusakov [1], *Junior Researcher, rusakov.msk@yandex.ru*
A.A. Genov [2], *Dr.Sc. (Engineering), Professor, Senior Researcher, kt-mati@mail.ru*
S.Sh. Khil [3], *Ph.D. (Engineering), Associate Professor, avs57@mail.ru*

[1] *V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow,117997, Russian Federation*
[2] *Center of Visualization and Satellite Information Technologies SRISA, Moscow, 117218, Russian Federation*
[3] *Moscow Aviation Institute (National Research University), Moscow, 125993, Russian Federation*

Abstract. Nowadays, the problem of providing a high security level of mobile devices, such as smartphones and tablets, through biometric approaches is particularly relevant. The paper proposes an anti-spoofing method for a limited number of images. The authors investigate the detection of spoof attacks using printed photos and screens of mobile devices and monitors. They show the relevance of the research and current unresolved problems.

The paper considers the structure of a typical Liveness system consisting of a source (camera), preprocessing of received images, face detection, feature module and a classifier. The study shows that pre-processing of the obtained images is one of the most important parts of the system, since the signs of spoof attacks are difficult to detect due to the extensiveness of augmentations.

There is a small review of modern architectures of convolutional neural networks (called encoders in the current architecture terminology). It is also shown that the linear output of convolutional neural networks can be used as an input for LSTM-type recurrent neural networks. It is noted that the best face detection algorithm in the current architecture is MMOD method.

The resulting Liveness system architecture is presented as a combined approach consisting of two components: a convolutional neural network receiving embeddings from each frame and a LSTM recurrent neural network that uses these insertions on the input and learns to remember sequences of certain frames and their characteristics.

The authors present experimental results showing the accuracy indicators of the current developments in this field when computing power requirements are low. This methodology allows determining fake photos by several frames.

Keywords: anti-spoofing, liveness, face recognition, convolutional neural networks, recurrent neural networks.

Acknowledgements. *The publication is a part of the state assignment of the SRISA RAS (fundamental research GP 14), topic No. 0065-2019-0001 (AAAA-A19-119011790077-1).*

References

1. Bhaganagare B.B., Harale A.D. Iris as biometrics for security system. *Proc. ICECCT, Coimbatore*. 2017, pp. 1–7. DOI: 10.1109/ICECCT.2017.8117952.

2. Ivanov V.I., Baras J.S. Authentication of fingerprint scanners. *Proc. IEEE ICASSP*. Prague, 2011, pp. 1912–1915.

3. Pan G., Sun L., Wu Z., Lao S. Eyeblink-based antispoofing in face recognition from a generic Webcamera. *Proc. IEEE 11th ICCV'07*. Rio de Janeiro, Brazil, 2007, pp. 14–20. DOI: 10.1109/ICCV.2007.4409068.

4. Kollreider K., Fronthaler H., Faraj M.I., Bigun J. Real-time face detection and motion analysis with application in 'liveness' assessment. *Proc. IEEE Trans. Inf. Forensics Security*. 2007, vol. 2, no. 3, pp. 548–558. DOI: 10.1109/TIFS.2007.902037.

5. Sun L., Pan G., Wu Z., Lao S. Blinking-based live face detection using conditional random fields. *Proc. Intern. Conf. Adv. Biometrics*. Seoul, Korea, 2007, pp. 252–260.

6. Anjos A., Chakka M.M., Marcel S. Motion-based counter-measures to photo attacks in face recognition. *IET Biometrics*. 2014, vol. 3, no. 3, pp. 147–158.

7. Charnes A., Cooper W.W., Rhodes E. Measuring the efficiency of decision-making units. *Europ. J. Operation Research.* 1978, vol. 2, pp. 429–444.

8. Kim Y., Na J., Yoon S., Yi J. Masked fake face detection using radiance measurements. *J. Opt. Soc. Amer. A*. 2009, vol. 26, no. 4, pp. 760–766.

9. Kim W., Suh S., Han J. Face liveness detection from a single image via diffusion speed model. *Proc. IEEE Trans. Image Process*. 2015, vol. 24, no. 8, pp. 2456–2465. DOI: 10.1109/TIP.2015.2422574.

10. Efremov I.A., Mamrosenko K.A., Reshetnikov V.N. Methods of developing graphics subsystem drivers. *Software & Systems.* 2018, no. 3, pp. 425–429. DOI: 10.15827/0236-235X.123.425-429 (in Russ.).

11. Liu Y., Jourabloo A., Liu X. Learning deep models for face anti-spoofing: Binary or auxiliary supervision. *Proc. CVPR, IEEE*. 2018. Available at: https://ieeexplore.ieee.org/document/8578146/authors#authors (accessed July 20, 2019). DOI: 10.1109/CVPR.2018.00048.

12. King D.E. *Max-Margin Object Detection.* 2015. Available at: https://arxiv.org/abs/1502.00046 (accessed July 20, 2019).

13. Hochreiter S., Schmidhuber J. Long short-term memory. *Neural Computation*. 1997, vol. 9, no. 8, pp. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.

14. Budylsky D.V. GRU and LSTM: modern recurrent neural networks. *Young Scientist.* 2015, no. 15, pp. 51–54 (in Russ.).

Для цитирования

Русаков К.Д., Генов А.А., Хиль С.Ш. Методика решения задачи антиспуфинга по ограниченному количеству фотографий // Программные продукты и системы. 2020. Т. 33. № 1. С. 054–060. DOI: 10.15827/0236-235X.129.054-060.

For citation

Rusakov K.D., Genov A.A., Shil S.Sh. An anti-spoofing methodology for a limited number of photos. *Software & Systems.* 2020, vol. 33, no. 1, pp. 054–060 (in Russ.). DOI: 10.15827/0236-235X.129.054-060.